

# **METHOD AND SYSTEM FOR CLINICAL KNOWLEDGE MANAGEMENT**

## **RELATED PROVISIONAL APPLICATION**

5           The present application claims the benefit of U.S. Provisional Patent  
Application entitled Method And System For Clinical Knowledge Management, filed  
on March 23, 2000, application number 60/191,524 and US. Provisional Patent  
Application entitled Method And System For Clinical Knowledge Management, filed  
on March 24, 2000, application number 60/191,967.

10

## **FIELD OF THE INVENTION**

          The present invention relates to computer networks and information systems.  
In particular, the present invention pertains to a computer based system and method  
for the analysis and evaluation of medical conclusions.

15

## **BACKGROUND INFORMATION**

          Medicine is one of the most data intensive, but least automated industries.  
The present state of medical informatics is limited to demographics. Clinical data is  
expressed with rudimentary information available from ICD and CPT codes. Current  
20   technology relies upon representation of clinical information in text form using  
phrases. Expert systems technology has also been developed in order to provide  
automated computer support to the practice of medicine. However, these systems  
generally do not provide meaningful analysis of medical data based upon discrete  
temporal segments.

25           In particular, current approaches are unable to convey generalities or nuances  
of data that can easily be manipulated and analyzed via information processing  
devices such as computers. Each segment in a medical history of a patient relating to  
a claim necessarily involves one or more clinical conclusions, which inform a  
treatment plan, medical tests ordered, diagnoses and follow-up procedures. However,  
30   known expert systems for providing medical analysis do not allow the evaluation of  
medical clinical conclusions relating to particular segments of a medical history.

          In order to provide meaningful representation and analysis of medical data, it  
is necessary to develop a codification system that emulates the methodology  
employed by doctors and other health care professionals.

35

EL 17995491405

## SUMMARY OF THE INVENTION

The present invention provides a method and system for evaluation of a medical conclusion such as a diagnosis. The present invention may be applied either retrospectively to evaluate a prior medical conclusion or prospectively to evaluate one  
5 or more current medical conclusions.

According to one embodiment, a medical analysis site is coupled to an information network such as the Internet. Clients including medical payors, medical providers and employers interact with the medical analysis site via the information network. The medical analysis site receives input data from clients either in the form  
10 of raw medical documents, which are further processed at the medical analysis site or via a data input interface from an operator.

The medical analysis site includes a front-end subsystem (e.g., a World Wide Web server) for providing a graphical user interface ("GUI") for clients to interact with the site, a core engine, which performs at least one process for analyzing and  
15 evaluating medical conclusions and a patient record database, which stores normalized medical data relating to medical histories of patients. The medical analysis site stores a predefined set of medical essential elements to represent and normalize events in a medical history including symptom reports, treatments, clinical conclusions, laboratory tests, chronological factors, demographic factors, etc.

According to one embodiment of the present invention, the core engine  
20 includes a processor and a relational database, which further includes a medical essential element database, a medical phrase database, a chronological rules database and a medical knowledge rules database. The medical phrase database maps at least one medical phrase to an essential element. The medical knowledge rules database  
25 maps each essential element to at least one clinical conclusion using a membership confidence function and a criterion impact parameter. The membership confidence function relates a degree of confidence that a particular essential element points to a particular clinical conclusion as a function of a medical assessment. The criterion  
30 impact parameter expresses an importance weight to be assigned a particular essential element with respect to a particular clinical conclusion. The chronological rules database stores information, which is used to segment medical data into discrete segments for further processing.

The medical analysis site also includes a patient record database, which stores normalized patient claim medical data, which is processed by the core engine.

The core engine performs an encoding process, a data structuring/sequencing process and a clinical conclusion/reporting process. During the encoding process, the core engine maps a set of medical data relating to one or more medical claims to a set of essential elements. In particular, according to one embodiment, the medical analysis site receives medical data relating to patient claims. The medical data may be submitted in the form of raw medical documents in which case OCR ("Optical Character Recognition") technology is employed to render the data in a computer readable form. In an alternative embodiment, an operator may directly submit medical data to the medical analysis site via input forms generated by the front-end subsystem. The core engine parses received medical data into discrete phrases, which are compared with data stored in the medical essential element database. If a matching phrase is found, a corresponding essential element record is instantiated in the patient record database. The core engine then segments medical elements stored in the patient record database relating to a particular medical claim into discrete temporal segments using chronological rules stored in the chronological rules database.

The core engine also performs a clinical conclusion analysis and reporting process to evaluate one or more clinical conclusions with respect to the temporal segments determined in the sequencing process. In particular, data gleaned from a patient's medical experience is compared with a known "best case" scenario for each specific clinical conclusion. The results of the clinical conclusion and analysis process are then reported in graphical form for review.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1a depicts a relationship between medical data and a number of essential medical elements according to one embodiment of the present invention.

FIG. 1b depicts an exemplary hierarchical structure for representing a relationship between essential elements according to one embodiment of the present invention.

FIG. 1c depicts a relationship between a number of utility groups and classes according to one embodiment of the present invention.

FIG. 1d depicts a schematic of a process for evaluating medical clinical conclusions according to one embodiment of the present invention.

FIG. 2 depicts a network architecture illustrating the relationship between a medical analysis site and various clients.

FIG. 3a depicts the architecture of a core engine relational database according to one embodiment of the present invention.

5        FIG. 4a depicts a data structure for storing an essential element record according to one embodiment of the present invention.

FIG. 4b depicts a data structure for storing a phrase record in a phrase database according to one embodiment of the present invention.

10       FIG. 4c depicts a data structure for storing a chronological rule record according to one embodiment of the present invention.

FIG. 4d depicts a data structure for storing a medical rule record according to one embodiment of the present invention.

FIG. 5a depicts the hierarchical structure of a patient record database according to one embodiment of the present invention.

15       FIG. 5b depicts a data structure for representing an exemplary patient record according to one embodiment of the present invention

FIG. 5c depicts a data structure for representing a claim record according to one embodiment of the present invention.

20       FIG. 5d depicts a data structure for representing a temporal segment according to one embodiment of the present invention.

FIG. 5e depicts a data structure for representing an event according to one embodiment of the present invention.

FIG. 6a is a flowchart that depicts a series of steps for performing analysis of medical data according to one embodiment of the present invention

25       FIG. 6b is a flowchart of steps depicting document parsing/phrase matching/encoding.

FIG. 6c is a flowchart of steps depicting a sequencing process according to one embodiment of the present invention.

30       FIG. 7 is a flowchart of steps for generating a clinical conclusion report according to one embodiment of the present invention

FIG. 8a shows a sample report according to one embodiment of the present invention.

FIG. 8b shows an exemplary report including various sample data points according to one embodiment of the present invention.

FIG. 9a illustrates an exemplary fuzzy logic membership function relating a measurement value for an essential element with a confidence parameter according to one embodiment of the present invention.

FIG. 9b illustrates the evaluation of a clinical conclusion with respect to a  
5 number of essential elements according to one embodiment of the present invention.

### **DETAILED DESCRIPTION**

FIG. 1a depicts a relationship between medical data and a number of essential  
10 medical elements according to one embodiment of the present invention. Medical data typically includes a myriad of discrete data elements (herein referred to as “essential elements”) such as patient symptoms, test reports, clinical conclusions, etc, which hold significance in evaluating or performing a diagnosis. Typically, these discrete data elements are distributed in patient records as a function of test results, evaluations, examination, etc. FIG. 1a depicts exemplary medical records 101a and  
15 101b, which each medical record includes a plurality of medical phrases 105a-105p. Medical phrases 105 may be gleaned from medical documents 101 via optical character recognition (“OCR”) technology or via manual text entry via an operator. Medical phrases may then be identified with particular essential elements in order to normalize the medical data. For example, different doctors may utilize slightly  
20 different terminology to refer to a common symptom. By identifying particular discrete data elements with essential elements, variations in phraseology or terminology may be eliminated and patient data normalized to a predefined set of medical elements.

FIG. 1b depicts an exemplary hierarchical structure for representing a  
25 relationship between essential elements according to one embodiment of the present invention. As shown in FIG. 1b, essential elements 107(1)-107(N) are included in category 106(1), categories 106(1)-106(N) are included in class 104(1) and classes 104(1)-104(N) are included in utility group 109. Note that FIG. 1b depicts a single utility group 109. Any arbitrary number of utility groups 109, each of which utilizes a  
30 unique relationship for the representation of essential elements may be defined. For example, FIG. 1c depicts a relationship between a number of utility groups 109 and classes according to one embodiment of the present invention. As shown in FIG. 1c, a non-clinical data utility group 109 includes, among other classes, symptoms class 107, predisposing factors class 107, treatment class 107, diagnostic tests class 107,

etc. Thus, for example, symptoms class 107 might include a number of categories 106 and/or essential elements 107 that relate to various symptoms. Chronological segment utility group 109 includes various classes relating to temporal markers in a medical history. For example, chronological segments may include various events in a medical history that represent natural segmenting points such as operative procedures, office visits, etc.

FIG. 1d depicts a schematic of a process for evaluating medical clinical conclusions according to one embodiment of the present invention. As described in detail below, according to one embodiment, the present invention provides a system, which receives a plurality of medical phrases relating to a medical history of a patient. The system provides output of an overall level of confidence parameter relating to various clinical conclusions derived in the medical history. According to one embodiment of the present invention, a database of essential medical elements is maintained, which is utilized to normalize medical phrases either gleaned from medical records or provided via operator input. Medical essential elements may relate to clinical data such as etiology, predisposing factors, inciting events, etc., demographic information, diagnosis etc. For example, a patient may have an initial consultation with a doctor relating to a particular medical condition. During the initial consultation, various significant medical events may occur such as laboratory tests, administration of various treatments,

In a document parsing/phrase matching phase 115, medical data is received, parsed and normalized according to a pre-defined set of medical essential elements. According to one embodiment, as shown in FIG. 1b medical data is received in the form of raw medical records 101. Medical record 101 is parsed to locate relevant medical phrases 105 utilizing phrase database 114. The phrases are then resolved into essential medical elements 107 utilizing essential element database 112.

In a data structuring/sequencing phase 120, normalized medical data is segmented into temporal frames 150(a)-150(d) based upon sequencing rules 117. In a clinical conclusion analysis and reporting phase 125, the temporal frames 150-150(d) are analyzed with respect to one or more clinical conclusions and reports are generated. Note, that FIG. 1d depicts only 4 temporal frames (150a-150d) but the present invention is compatible with the generation of any number of temporal frames. The details of each of these processes are discussed in detail below.

FIG. 2a depicts a network architecture illustrating a relationship between a medical analysis site and various clients. In particular, FIG. 2a illustrates a relationship between payor client 205a, provider client 205b, employer client 205c, medical analysis site 219 and service bureau 275.

5 Medical analysis site 219 is associated with one or more IP ("Internet Protocol") addresses, which correspond to a unique destination address on Internet 214. Client 205c is typically associated with a dynamically assigned unique IP address, which is assigned upon dial-up by ISP 220. The unique IP addresses of client 205c and medical analysis site 219 permit transmission of data between client  
10 205c and medical analysis site 219 via Internet 214. In particular, personal computer 212 packetizes data destined for medical analysis site 219 with a header, which includes the IP address of medical analysis site 219. This data is read and recognized by various routers 235 coupled to Internet 214, which route data packets containing the IP address of medical analysis site 219 to medical analysis site 219. Similarly,  
15 medical analysis site packetizes data destined for client 205c with the dynamically assigned IP address of client 205c, which is used to route data to client 205c.

As shown in FIG. 2, client 205c uses browser software 287 running as a separate process on personal computer 212 to transmit electronic signals representing information to medical analysis site 219 and to receive electronic signals from  
20 medical analysis site 219. Browser software 287 permits navigation between various file servers connected to Internet 214, including Web server 240 at medical analysis site 219. Browser software 287 also provides functionality for rendering of files distributed on the Internet (i.e., through plug-ins or Active X controls), which are displayed on display device 285.

25 Personal computer 212 communicates with Internet service provider ("ISP") 220 through a dial-up connection using modem 215, through POTS line 217 to a central office (not shown) and the public switched telephone network ("PSTN") (not shown). Typically, the transmission path from modem 215 through POTS line 117 to the central office is analog. At the central office, signals are sampled for digital  
30 transmission through the PSTN to ISP 220. Due to the analog nature of the transmission path from modem 215 through POTS line 217 to the central office, modem 215 performs modulation of digital signals generated by personal computer 212 onto an analog carrier signal for transmission to the central office. Modem 215

also performs demodulation of signals received over local lines (e.g., 217) from the central office, extracting digital byte codes from a modulated analog carrier.

ISP 220 is coupled to the PSTN through modem bank 221, which converts an analog modulated signal to a digital signal. Digital IP packets are then transmitted via Internet 114 and various routers (not shown) to Web server 240. IP packets are also transmitted in the reverse direction from Web server 240 to personal computer 212 via a path through Internet 214 and other routers, through ISP 220, modem bank 221, PSTN, central office and modem 215.

Network protocols and the network protocol stack in relationship to the OSI ("Open System Interconnection") for communication over Internet 214 and Web are well known. According to one embodiment of the present invention, browser software 287 uses HTTP to retrieve a particular hypertext Web page assigned a particular URL on the Web. Each HTTP interaction consists of one ASCII ("American Standard Code for Information Interchange") request followed by one RFC ("Request For Comments") 822 MIME ("Multipurpose Internet Mail Extensions") response. The HTTP connection is conducted over transmission control protocol (TCP) network layer. The TCP layer provides for a reliable stream-oriented connection between a server (e.g., 240) and personal computer 212 by resending data if necessary due to network overloads or malfunctions. Typically both personal computer 212 and Web server 240 run a TCP process that manages TCP streams and interfaces to an IP (Internetwork Protocol) layer. The TCP entity accepts user data streams from local processes, breaks them into segments and sends each piece as a separate IP datagram. On the other hand, when IP datagrams containing TCP data arrive at a machine, they are passed to the TCP process for reconstruction of the original byte streams.

The IP protocol at the network layer provides an addressing mode for sending packetized data from personal computer 212 to Web server 240 and vice versa. Point to point protocol ("PPP") at the data link layer primarily provides a framing method for sending data to the physical layer. PPP also allows IP addresses to be negotiated at connection time. The PPP protocol encodes the IP/TCP/HTTP packets and transmits them to modem 215 and the physical layer. The physical layer is the actual physical medium through which communications occur, e.g., twisted pair, coaxial cable, optical fiber, etc.

Although according to the embodiment described herein, HTTP communication between personal computer 112 and Web server 140 is conducted over a TCP/IP connection, other communication protocols are possible and this example is not intended to limit the claims appended hereto. For example, a UDP  
5 (“User Datagram Protocol”) implementation is possible.

It is to be understood that in general clients 205 may connect to Internet 114 using any potential medium whether it be a dedicated connection such as a cable modem, T1 line, DSL (“Digital Subscriber Line”) or a dial-up POTS connection. For example, in addition to client 205c, FIG. 2a also shows clients 205a and 205b. Client  
10 205b is coupled to Internet 214 via cable modem 268 and ISP 220. Client 205a, on the other hand, is a corporate client that is coupled to Internet 214 via T1 line 230b and router 235b. In this case, various node terminals (not shown) at client 205a share the bandwidth on T1 line 230b, wherein the bandwidth is distributed via Ethernet 274.

Service bureau 275 also communicates with medical analysis site 219 via  
15 Internet 214.

Medical analysis site 219 includes front-end subsystem 219, core engine 221 and patient record database 231. Front-end subsystem 219 provides a graphical user interface (“GUT”) for clients connecting to medical analysis site 219. In particular, front-end subsystem 219 includes web server 240a and HTML (“Hypertext Markup  
20 Language”) database 250a. Web server 240a serves HTML pages from HTML database 250a to clients 205 connecting with medical analysis site 219. Core engine 221 performs various backend processing functions at medical analysis site 219 related to the analysis of medical data as described in more detail below. Web server 240a is coupled to core engine server 240b at core engine 220, which is coupled to  
25 core engine relational database 250b. Patient record database 231 stores medical claim data relating to various patients in a normalized format as described in more detail below. Web server 240a is also coupled to medical data server 240c, which is coupled to medical data relational database 250.

FIG. 3 depicts the architecture of a core engine relational database according  
30 to one embodiment of the present invention. Core engine relational database 250b establishes a relational structure between essential elements 310, phrases 305, chronological rules 320, clarification rules 315 and medical knowledge rules 325.

FIG. 4a depicts a data structure for storing an essential element record according to one embodiment of the present invention. Each essential element record

405 includes class field 410, specificity field 412, category field 414, ID field 416, value field 418, value code field 420 and modifier field 422. Class field 410 stores a two-byte class that refers to the class of an essential element 107. Specificity field 412 stores an integer value representing a degree of specificity associated with an essential element. For example, an essential element "Operative Procedure" might be assigned a specificity value of 0, while an essential element "Endoscopic Carpal Tunnel Release – Chow Technique" might be assigned a specificity value 9 because the latter essential element engenders a higher degree of specificity with respect to particular clinical conclusions than the former. ID field 416 stores a 32-bit ID field derived from conventional code sources such as ICD or CPT. Value field 418 stores a binary value that represents whether an essential element is quantified with a numeric value (i.e., value code field 420). For example, some essential elements refer to parameters that may be measured during an examination etc. such as body weight etc. In this case, value field 418 would store a binary 'TRUE' value indicating that the essential element is quantified. If value field 418 is 'TRUE', value code field 420 stores either a floating-point value representing a quantified essential element value. Modifier field 422 stores a text string relating to other information related to the essential element.

FIG. 4b depicts a data structure for storing a phrase record in a phrase database according to one embodiment of the present invention. Each phrase record 407 includes an essential element ID 424 field and at least one phrase match string 426(1)-426(N). Essential element ID field 424 stores a unique 32-bit value of an essential element (i.e., from ID field 416). Phrase match string fields 426(1)-426(N) each store an ASCII ("American Standard Code For Information Interchange") string to be mapped to the essential element with ID stored in field 424.

FIG. 4c depicts a data structure for storing a chronological rule record according to one embodiment of the present invention. Each chronological rule record 409 includes act weight field 432 and scene weight field 414. Act weight field 432 stores a floating point value between 0-1 that represents the degree of confidence a particular event constitutes a change in temporal segment (e.g., an act). Scene weight field 434 stores a floating-point value between 0-1 that represents the degree of confidence that a particular event constitutes a change in a scene temporal segment.

FIG. 4d depicts a data structure for storing a medical rule record according to one embodiment of the present invention. Each medical rule record 411 includes

essential element ID field 436, conclusion field 438,  $\mu$  value field 442,  $\iota$  value field 444, value code field 446 and modifier field 448. Essential element ID field 436 stores the ID of an essential element associated with a clinical conclusion.

Conclusion field 438 stores a 32-bit integer value corresponding to a code for a  
5 particular clinical conclusion.  $\mu$  value field 442 stores a floating-point value between 0-1 that represents a degree of confidence that the essential element referenced in field 436 leads to the impression of the conclusion referenced in field 438.  $\iota$  value field 444 stores an integer value between 0-10 that represents an importance or impact that the essential element referenced in field 436 has on making the decision that the  
10 clinical conclusion referenced in conclusion ID field 440 is a member of a fuzzy set defined by that conclusion. Value code field 446 stores a 32-bit integer value that relates to a particular value to be associated with the essential element. According to one embodiment, this field may contain a pointer that references an analytical function that defines  $\mu$  as a function of an independent variable. Modifier field stores  
15 a text string that further describes the essential element referenced in field 436.

FIG. 5a depicts a hierarchical structure of a patient record database according to one embodiment of the present invention. At least one medical element 513 is associated with an event record 511. At least one event record 511 is associated with a temporal segment record 509. At least one temporal segment record 509 is  
20 associated with a claim record. At least one claim record 507 is associated with a patient record 505.

FIG. 5b depicts a data structure for representing an exemplary patient record according to one embodiment of the present invention. Each patient record 505 includes patient ID field 521, last name field 523, first name field 525, middle name  
25 field 527, date of birth field 529, street address field 531, city field 533, state field 535 and telephone field 537.

Patient ID field 521 stores a unique 32-bit integer value identifying a patient. Last name field 523, first name field 525, middle name field 527, date of birth field 529, street address field 531, city field 533, state field 535 and telephone field 537  
30 each store an ASCII string representing a last name, first name, middle name, date of birth, street address, city, state and telephone number of a patient respectively.

FIG. 5c depicts a data structure for representing a claim record according to one embodiment of the present invention. Each claim record 507 includes claim ID

field 539, patient ID field 540, claim number field 541, payor ID field 543 and adjuster ID field 545. Claim ID field 539 stores a unique 32-bit integer of a particular medical claim. Patient ID field 540 stores a patient ID (i.e., patient ID field 521) of the patient filing the claim. Claim number field 541 stores a unique 32-bit integer value representing a claim number. Payor ID field 543 and adjuster ID field 545 each respectively store 32-bit integer values relating to an ID of a payor and adjuster on the claim respectively.

FIG. 5d depicts a data structure for representing a temporal segment according to one embodiment of the present invention. According to one embodiment, each temporal segment is referred to as an act, which includes a set of events. For example, a temporal segment may be defined by a visit to a doctor's office, in which a variety of tests were performed, each comprising an element. As shown in FIG. 5d, each act record 509 includes act ID field 547, claim ID field 549, payor ID field 551 and adjuster ID field 553. Act ID field stores a unique 32-bit integer value defining an act. Claim ID stores the claim ID of a claim, which includes the act (i.e., claim ID field 539). Payor ID field 551 and adjuster ID field 553 each respectively store 32-bit integer values relating to the ID of a payor and adjuster on the claim respectively.

FIG. 5e depicts a data structure for representing an event according to one embodiment of the present invention. Each event record 511 includes event ID field 555, act ID field 557, event date field 559 and event type ID field 561. Event ID field stores a unique 32-bit integer corresponding to an event. Act ID field 557 stores an act ID number (i.e., field 547) corresponding to the event. Event date field 559 stores a date object variable representing a date upon which an event occurred. Event type ID field 561 stores a unique 32-bit integer value representing an event type.

FIG. 5f depicts a data structure for representing a medical element according to one embodiment of the present invention. An element record 513 is created for each medical phrase for which a corresponding essential element is found in essential element database 112. Each element record 513 includes element ID field 563, event ID field 565, essential element ID field 567, claim ID field 569, act ID field 571, date field 573 and event source ID field 575. Element ID field stores a unique 32-bit integer value generated for the element. Event ID field 565 stores an event ID of an event corresponding to the element (i.e., event ID field 555). Essential element ID field 567 stores an identifier of an essential element 107 corresponding to the element. Claim ID field 569 and act ID field 571 each respectively store an identifier of a

corresponding claim and act. Date field 573 stores a date object variable representing a date upon which the element occurred. Event source ID field 575 stores an ASCII string describing a source of the essential element.

FIG. 6a is a flowchart that depicts a series of steps for performing analysis of medical data according to one embodiment of the present invention. According to one embodiment, the steps depicted in FIG. 6a are carried out by core engine server 240b at medical analysis site 219. The process is initiated in step 605. Steps 610 and 615 correspond to document parsing/phrase matching/encoding step 115 depicted in FIG. 1d. In step 610, medical data is parsed into phrases and a lookup operation is performed using phrase database 114 to determine matching phrases. In step 615, matching phrases are encoded into essential elements 107 by instantiating a medical element record 513 located for each matching phrase.

Steps 617 and 619 correspond to data structuring/sequencing process 120 in FIG. 1d. In step 617, chronological sequencing of all medical element records 513 instantiated in step 615 is performed. In particular, as described in detail below, medical element records 513 are ordered chronologically using date field 573 of element record 513. In addition, medical elements are segmented into temporal segments (e.g., acts) utilizing sequencing rules database 117 described below.

Steps 621 and 623 correspond to clinical conclusion and reporting process 125. In step 621, conclusion analysis is performed upon one or more temporal segments with respect to one or more clinical conclusions. In step 623, the outcome of clinical analysis step 621 is reported. The process ends in step 625.

#### **Document Parsing/Phrase Matching Encoding Process**

FIG. 6b is a flowchart of steps depicting document parsing/phrase matching/encoding process 115 (i.e., steps 610 and 615 of FIG. 6a). The process depicted in FIG. 6b corresponds to the input of medical data relating to one medical claim and may be utilized either for automatic document parsing (e.g., using OCR) or for manual operator input of medical data. Practitioners skilled in the art will recognize that the process depicted in FIG. 6b may be elaborated or modified to receive data for multiple claims or multiple patients.

The process is initiated in step 629. In step 632a new claim record 507 is instantiated. In step 637, a next phrase is retrieved. According to one embodiment, the next phrase is retrieved via an OCR process from a medical document. In the

alternative, if the input is provided manually, the next phrase is received from a human operator interacting with medical analysis site 219. In step 639, phrase database 114 is searched to locate a matching phrase. In step 641, it is determined whether the phrase is located in phrase database 114. If not ('no' branch of step 641), the next phrase is considered in step 637. If the phrase is located in phrase database ('yes' branch of step 641), in step 643, a new medical element record 513 is instantiated. In particular, a unique element ID is generated, which is used to populate element ID field 563. The essential element ID corresponding to the matched phrase is stored in essential element ID field 567 and the current claim ID is stored in claim ID field 569. The date pertaining to the element is stored in date field 573. In step 645, it is determined whether all phrases have been analyzed. If not ('no' branch of step 645, flow continues with step 637 and the next phrase is analyzed. If all phrases have been analyzed ('yes' branch of step 645), in step 651 a sequencing process is initiated.

#### **Data Structuring/Sequencing Process**

FIG. 6c is a flowchart of steps depicting a sequencing process according to one embodiment of the present invention. The process is initiated in step 652. In step 653, all medical element records 513 are chronologically sorted based upon date field 573. In step 654, a new act record 509 is instantiated and set as the current act record. The ID of the current claim is then stored in claim ID field 549. In step 656, a new event record 511 is instantiated and set as the current event. The current act ID is stored in act ID field 557. In step 657, it is determined whether any of the sequenced elements have not been analyzed. If not ('no' branch of step 657), the process ends in step 665.

If there are remaining elements ('yes' branch of step 657), in step 659 it is determined whether the next element is an event class element (i.e. is included in the event class). If so ('yes' branch of step 659), a new event record is instantiated in step 661 and set as the current event. In step 663, it is determined whether a new act should be established based upon chronological rules. In particular, according to one embodiment of the present invention, a lookup process is performed using the chronological rules database to locate the event corresponding to the current event and the corresponding rule is applied to determine whether a new act should be

established. According to one embodiment, for example, the  $\mu$  value for the current event is retrieved from the chronological rules database. If the  $\mu$  value exceeds a certain value such as 0.7, a new act is established unless a new act occurred less than fourteen days prior. If the chronological rules dictate that a new act should be established ('yes' branch of step 663), in step 662 a new act record is instantiated and the current event is assigned to the new act record. Flow continues with step 657. If a new act is not necessitated ('no' branch of step 663), flow continues with step 657. If the current element is not an event ('no' branch of step 659), in step 660, the element event ID is set to be equal to the current event ID. Flow continues with step 657.

10

### **Clinical Conclusion Analysis And Reporting**

After data structuring and sequencing, core engine 221 performs a clinical conclusion analysis and reporting process. Normalized patient data relating to a particular claim stored in patient record database 231 is analyzed and compared with a best-case scenario with respect to each clinical conclusion. According to one embodiment, the following methodology is employed:

15

Let  $EE = \{y : y \text{ is an essential element}\}$

Where  $y \in EE$ , let  $\max(f(y)) = y_{\max}$ , where  $f(y_{\max})$  is maximum for all  $y \in EE$

Let  $\mu(y, cc) = \text{membership confidence value of } y \text{ w.r.t. clinical conclusion } cc$ , where  $y \in EE$

Let  $\iota(y, cc) = \text{criterion impact value of } y \text{ w.r.t. clinical conclusion } cc$ , where  $y \in EE$

20

Let  $\alpha(y, cc) = \mu(y, cc)\iota(y, cc)$

Let  $BC_{cc} = \{y \in EE : y = \max(\mu(y, cc)) \text{ and } y = \max(\iota(y, CC))\}$

Let  $P_T = \{z : z \in EE \text{ and } z \text{ is associated with a patient element belonging to temporal segment } T\}$

Let  $CA = \{x : x \text{ is a category}\}$

Let  $CL = \{x : x \text{ is a class}\}$

Let  $Ca_n = \{z : z \in EE \text{ and } z \text{ belongs to category } n\}$

Let  $Cl_n = \{w : w \in CA \text{ and } w \text{ belongs to class } n\}$

25

$$\mu(Ca_n) = \max(\mu(z)) \text{ where } z \in Ca_n$$

$$i(Ca_n) = \sum_{k=1}^N i(z_k, cc_i)(\mu(z_k, cc) - 0.5), \text{ where } z_k \in Ca_n$$

$$\alpha(Ca_n) = \mu(Ca_n)i(Ca_n)$$

$$\lambda(Ca_n) = \alpha_p(Ca_n) / \alpha_b(Ca_n)$$

$$\mu(Cl_n) = \max(\mu(z)) \text{ where } z \in Cl_n$$

$$i(Cl_n) = \sum_{k=1}^N i(z_k)(\mu(z_k) - 0.5), \text{ where } z_k \in Cl_n$$

$$\alpha(Cl_n) = \mu(Cl_n)i(Cl_n)$$

$$\lambda(Cl_n) = \alpha_p(Cl_n) / \alpha_b(Cl_n)$$

$$M = \max(\mu(z)) \text{ where } z \in Cl_n$$

$$I = \sum_{k=1}^N i(z_k)(\mu(z_k) - 0.5), \text{ where } z_k \in CL$$

$$A = MI$$

$$\Lambda = A_p / A_b$$

FIG. 7 is a flowchart of steps for generating a clinical conclusion report

- 5 according to one embodiment of the present invention. It is assumed that it is desired to evaluate a particular temporal segment (e.g., an act) with respect to a particular clinical conclusion. Practitioners skilled in the art will recognize that the following description may be easily extended to evaluate multiple clinical conclusions over multiple temporal segments. The process is initiated in step 705. In step 710,
- 10 parameters  $i_b(Ca_n)$ ,  $\mu_b(Ca_n)$  and  $\alpha_b(Ca_n)$  are calculated for each category of  $BC_{cc}$ . In step 720, parameters  $i_p(Ca_n)$ ,  $\mu_p(Ca_n)$  and  $\alpha_p(Ca_n)$  are calculated for each category of  $P_T$ . In step 730,  $\lambda(Ca_n) = \alpha_p(Ca_n) / \alpha_b(Ca_n)$  is calculated. In step 735, parameters  $i_b(Cl_n)$ ,  $\mu_b(Cl_n)$  and  $\alpha_b(Cl_n)$  are calculated for each class of  $BC_{cc}$ . In step 740,
- 15 parameters  $i_p(Cl_n)$ ,  $\mu_p(Cl_n)$  and  $\alpha_p(Cl_n)$  are calculated for each class of  $P_T$ . In step 750, parameters  $M_p$ ,  $I_p$  and  $A_p$  are calculated. In step 760, parameters  $M_b$ ,  $I_b$  and  $A_b$  are calculated. In step 770,  $\Lambda = A_p / A_b$  is calculated. In step 780, a report is generated. The process ends in step 790.
- According to one embodiment, the following pseudo-code, describes a process for a
- 20 total criteria impact (I), a total membership confidence (M) and an overall level of confidence ratio  $\Lambda$ .

```

struct confidence_impact
{
float total_criteria_impact_best_case;
float total_membership_confidence_best_case;
5 float total_criteria_impact_patient;
float total_membership_confidence_patient;
}
confidence_impact* Conclusion_Analysis(eeset * elements, conclusion
conclusion_code)
10
{
Sort elements into categories;
Calculate  $\mu_b(Ca_n), \iota_b(Ca_n), \alpha_b(Ca_n)$ ;
Calculate  $\mu_p(Ca_n), \iota_p(Ca_n), \alpha_p(Ca_n)$ ;
15 Calculate  $\lambda(Ca_n)$ ;
Calculate  $\mu_b(Cl_n), \iota_b(Cl_n), \alpha_b(Cl_n)$ ;
Calculate  $\mu_p(Cl_n), \iota_p(Cl_n), \alpha_p(Cl_n)$ 
Calculate  $\lambda(Cl_n)$ ;
Calculate  $M_b, I_b, A_b$ ;
20 Calculate  $M_p, I_p, A_p$ ;
Calculate  $\Delta$ ;
Return confidence_impact;
}

```

FIG. 8a shows a sample report according to one embodiment of the present invention. FIG. 8 shows patient conclusion analysis result 805 compared with best-case conclusion analysis 803 for a single clinical conclusion with respect to a temporal segment. The area of patient conclusion analysis result 820= $M(b)$  830 multiplied by  $I(b)$  815. Similarly, the area of best-case conclusion analysis result 840= $M(p)$  860 multiplied by  $I(p)$  845.  $\Delta$  represents the ratio between the areas of the actual patient data and the best-case scenario =  $A(p)/A(b)$ .

FIG. 8b shows an exemplary report including various sample data points according to one embodiment of the present invention.

According to one embodiment of the present invention, value code field 446 of each medical rule record 411 stores a pointer that reference a function that defines  $\mu$  (confidence parameter) as a function of an independent variable. According to one embodiment, the function may be a fuzzy logic membership function. For example, a particular essential element 107 may be associated with a measurement value such that the confidence parameter  $\mu$  is determined as a function of the measurement value. Typically, the functional relation maps a certain measurement associated with an essential element 107 to a particular confidence parameter  $\mu$  for a particular clinical conclusion being considered.

FIG. 9a illustrates an exemplary fuzzy logic membership function relating a measurement value for an essential element with a confidence parameter according to one embodiment of the present invention. As shown in FIG. 9a, for a particular clinical conclusion and essential element 107, fuzzy logic membership function 910 maps a particular measurement value (e.g., 915) to a particular confidence parameter  $\mu$  (e.g., 917).

FIG. 9b illustrates the evaluation of a clinical conclusion with respect to a number of essential elements according to one embodiment of the present invention. It is assumed that a plurality of essential elements has been determined from raw medical data and corresponding medical rule records have been generated. Further, it is assumed that the determined essential elements 107 are associated with a measurement value 915 such that the confidence parameter  $\mu$  is determined as a function of the measurement value 915. As shown in FIG. 9b, each relevant essential element 107(1)-107(N) is considered by determining the membership confidence value  $\mu_1 - \mu_N$  utilizing corresponding function 910(1)-910(N). FIG. 9b illustrates one particular method for determining an overall level of confidence parameter. An overall level of confidence parameter is determined according to the following linear

combination of functions  $f(\mu_j, t_j)$ . 
$$M = \sum_{j=1}^N f(\mu_j, t_j)$$
. Note that functions

$f(\mu_j, t_j)$  may be any function such as a simple product weighting. Also, according to alternative embodiments,  $\Lambda$  may be determined by non-linear methods.